METHOD AND SYSTEM FOR THROTTLING I/O REQUEST SERVICING ON BEHALF OF AN I/O REQUEST GENERATOR TO PREVENT MONOPOLIZATION OF A STORAGE DEVICE BY THE I/O REQUEST GENERATOR

Abstract Of The Disclosure

A method and system for fairly distributing servicing of I/O requests by a storage device among remote computers that contract for a specified maximum rate of I/O request servicing by the storage device. The remote computers also specify whether the storage device should employ a simple throttling technique or a sliding window throttling technique to limit the rate of I/O request servicing to the contracted-for rate of I/O request servicing. In the simple throttling technique, the storage device services up to some maximum number of I/O requests per unit time for the remote computer, regardless of the overall time-averaged rate of I/O request servicing provided to the remote computer. In the sliding window throttling technique, the storage device maintains an approximate, periodically re-calculated instantaneous rate of I/O request servicing calculated based on recent I/O request servicing by the storage device on behalf of the remote computer. Intervals between servicing of I/O requests are lengthened or shortened in order to continuously adjust the instantaneous rate of I/O request servicing towards the contracted-for rate of I/O request servicing.